

D.C. Crouch



RAW SEQUENCE LISTING

DATE: 02/21/2002

PATENT APPLICATION: US/09/143,155

TIME: 09:52:14

Input Set : N:\Crf3\RULE60\09143155.txt

Output Set: N:\CRF3\02212002\I143155.raw

SEQUENCE LISTING

4 (1) GENERAL INFORMATION:

C--> 6 (i) APPLICANT: DiTullio, Paul A.; Meade, Harry; Cole, Edward S.

8 (ii) TITLE OF INVENTION: TRANSGENETICALLY PRODUCED ANTITHROMBIN III

10 (iii) NUMBER OF SEQUENCES: 2

12 (iv) CORRESPONDENCE ADDRESS:

13 (A) ADDRESSEE: LAHIVE & COCKFIELD, LLP

14 (B) STREET: 28 State Street

15 (C) CITY: Boston

16 (D) STATE: Massachusetts

17 (E) COUNTRY: USA

18 (F) ZIP: 02109

20 (v) COMPUTER READABLE FORM:

21 (A) MEDIUM TYPE: Floppy disk

22 (B) COMPUTER: IBM PC compatible

23 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

24 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

26 (vi) CURRENT APPLICATION DATA:

C--> 27 (A) APPLICATION NUMBER: US/09/143,155

C--> 28 (B) FILING DATE: 28-Aug-1998

29 (C) CLASSIFICATION:

31 (vii) PRIOR APPLICATION DATA:

32 (A) APPLICATION NUMBER: 08/391,743

33 (B) FILING DATE:

35 (viii) ATTORNEY/AGENT INFORMATION:

36 (A) NAME: Myers, Paul Louis

37 (B) REGISTRATION NUMBER: 35,965

38 (C) REFERENCE/DOCKET NUMBER: TCI-045

40 (ix) TELECOMMUNICATION INFORMATION:

41 (A) TELEPHONE: (617)227-7400

42 (B) TELEFAX: (617)742-4214

44 (2) INFORMATION FOR SEQ ID NO: 1:

46 (i) SEQUENCE CHARACTERISTICS:

47 (A) LENGTH: 3515 base pairs

48 (B) TYPE: nucleic acid

49 (C) STRANDEDNESS: single

50 (D) TOPOLOGY: linear

52 (ii) MOLECULE TYPE: DNA (genomic)

54 (iii) HYPOTHETICAL: NO

56 (iv) ANTI-SENSE: YES

58 (v) FRAGMENT TYPE: N-terminal

62 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

64 AGAATTTACC CCAAGATCTC AAAGACCCAC TGAATACTAA AGAGACCTCA TTGTGGTTAC

ENTERED

60

RAW SEQUENCE LISTING

DATE: 02/21/2002

PATENT APPLICATION: US/09/143,155

TIME: 09:52:14

Input Set : N:\Cr3\RULE60\09143155.txt

Output Set: N:\CRF3\02212002\I143155.raw

66	AATAATTTGG	GGACTGGGCC	AAAACCTCCG	TGCATCCCAG	CCAAGATCTG	TAGCTACTGG	120
68	ACAATTTTCAT	TTCCTTTATC	AGATTGTGAG	TTATTCCTGT	TAAAATGCTC	CCCAGAATTT	180
70	CTGGGGACAG	AAAAATAGGA	AGAATTCATT	TCCTAATCAT	GCAGATTTCT	AGGAATTCAA	240
72	ATCCACTGTT	GGTTTATTT	CAAACCACAA	AATTAGCATG	CCATTAAATA	CTATATATAA	300
74	ACAGCCACTA	AATCAGATCA	TTATCCATTC	AGCTTCTCCT	TCACTTCTTC	TCCTCTACTT	360
76	TGGAaaaaag	GTAAGAATCT	CAGATATAAT	TTCAGTGTAT	CTGCTACTCA	TCTTTATTTT	420
78	GGACTAGGTT	AAAATGTAGA	AAGAACATAA	TTGCTTAAAA	TAGATCTTAA	AAATAAGGGT	480
80	GTTTAAGATA	AGGTTTACAC	TATTTTCAGC	AGATATGTTA	AAAAATAGAA	GTGACTATAA	540
82	AGACTTGATA	AAAATTATAG	GTGACTGCAA	TTTTTGCCAT	GAGGTTTGCA	GGATCTTGGT	600
84	TCCCTGACCA	GGGATCAAAC	CTGCACTCCC	CTGGAAGCAT	GGAGTCTTGG	ACATTTGTAT	660
86	TATACACTAT	CTTTGGTTCC	TTTTAAAGGG	AAGTAATTTT	ACTTAAATAA	GAAAATAGAT	720
88	TGACAAGTAA	TACGCTGTTT	CCTCATCTTC	CCATTACACAG	GAATCGAGAG	CCATGAAGGT	780
90	CCTCATCCTT	GCCTGTCTGG	TGGCTCTGGC	CATTGCAAGA	ACAAAGAGGT	AAATACAGAA	840
92	AAAATGTTGA	AATAATAGAC	TAGTACTGTC	TGCCTATGTG	TAGAAATCAC	ATTACCAACA	900
94	TCATAAATGT	ATAAATAATG	CACAATCTCA	GATTTATTTT	TTAATGCTAA	GAAAGTCATT	960
96	TATGTTTCATC	CACTATCTCA	ACAGTATCCT	ATAGGACCAC	AACTCTGGGT	CAAGTGCTTT	1020
98	CTATAGTATT	GTACCATCTG	TACCATCAAT	TCCTAAAGAA	AAAGGAAAAG	AAACCAATAA	1080
100	GCAACAGACC	AACAAGAAGG	AACACAGACA	AGAACAAAAA	ATGAGTAATA	TTGTACAAAT	1140
102	ACAATTGCAC	GCTGCAGGAA	CTAAAGTGTT	TTTTTTTCTC	TCTCTTTTTA	GCAGGAAGAA	1200
104	CTCAATGTAG	TCGGTGAGGT	AAGATGATTT	TTATACAAAG	AAAAAAATTA	ATTTAACTGT	1260
106	AAAATAGTAA	CAGACTCTGA	TGATCTAGCA	GAAAACCTCAG	CTAATTGTCA	ATTTTTATTT	1320
108	TTCCTTTATA	GA CTGTGGAA	AGCCTTTTCA	AGCAGTGAGG	TAAGATAATG	TTCATTGAGA	1380
110	GGCAATTTCC	CAGATTTAGA	GCAATAAAAC	ACTGTATTAT	CTTTTTGTGT	TACATTAATT	1440
112	GGCAACCCAC	TCCAGTACTC	TTGACAAATA	TGAATTTTTT	TTTAAAGCTA	AACCTGATTT	1500
114	TATTTTTTATT	TTTCCAAAGG	AATCTATTAC	ACACATCAAT	AAGGTAAAAC	CTTCATATTT	1560
116	AAACGTATCA	TTTTTAAAAA	TTTCATGTTT	GATTTTTTATA	AACAGCATTT	CTTTATGTGT	1620
118	GATTTTTTTT	TTTACCAGAA	AATTGAGAAG	TTTCAAAGTG	AGGAACAACA	GCAAACAGAG	1680
120	GTAATTTGTT	TCACTATGAG	TATATTTTGA	GAAAGTATTAT	GAAACATAAC	ACATAAAAGA	1740
122	TTTATAATAA	TTATGTTTCT	TCTAAGAATG	GTAATATAAG	TGTCAGTGTA	AGAAATGAAA	1800
124	ACTTTGACAA	AATGAAAATA	TTTTAAAGAT	AGAAACACAT	TTTTAAACAC	ATAATCAAAT	1860
126	TTCAGAGTAT	AGAATAAATA	CCCAAGAATA	ACTACTGGTA	TATTCATTTT	ACTAATGGTA	1920
128	TACCTGGCTT	TAATAAATGC	ATATTAGTAG	GAACAATTCC	AGACTAGGGA	CTGTGATCCC	1980
130	CTTATTCTAA	TGATGGATAT	GCTGATGAAA	GACAGTAGGG	TGACAGTGTG	GCACTAATCC	2040
132	TAATAAATGG	AAGATTTTCT	TTCTCTCTCT	TCCTGAATT	ATGTTTTTAA	AAGAGGAGGA	2100
134	TAATTCATCA	TGAATAACAA	TTATAACTGG	ATTATGGACT	GCAAAGGCAT	TGGTTTTCTT	2160
136	TCTTTCCAGG	ATGAACTCCA	GGATAAAATC	CACCCCTTTG	CCCAGGCACA	AAGTCTCTAG	2220
138	TCTATCCCTT	CACTGGGCCC	ATCCCTAACA	GCCTCCCACA	AAACATCCTG	CTGCTCTTAC	2280
140	TCAAACCCCT	GTGGTGGTGC	CGCCTTTTCT	TCAGCCTGAA	ATAATGGGAG	TCCTTGCCAA	2340
142	AGTGAAGGAG	ACTATGGTTC	CTAAGCACAA	AGAAATGCCC	TTCCCTAAAT	ATCCAGTTTG	2400
144	AGCCCTTTAC	TGAAAGCCAG	AGCCTGACTC	TCCTGATGT	TGAAAAGCTG	CACCTTCTTT	2460
146	CTCTGCCTCT	GGTCCAGTCT	TGGATGCACC	AGCCTCCCCA	GCCTCTTTCT	CCAACCGTCA	2520
148	TTGTTTTCTC	CTCAGTCCGT	GCTGTCCCTT	TCTCAGCCCA	AAGTTCTGCC	TGTTCCCCAG	2580
150	AAAGCAGTGC	CCCAGAGAGA	TATGCCCATC	CAGGCCTTTT	TGCTGTACCA	GGAGCCTGTA	2640
152	CAATTGGTCC	TGTCCGGGGA	CCCTTCCCTA	TTCTTGTAAG	TCTAAATTTA	CTAACTGTGC	2700
154	TGTGGTTAAC	TTCTGATGTT	TGTATGATAT	TTGAGTAATT	AAGAGCCCTA	CAAAAAATC	2760
156	AATAATGAAT	GGTTCCAAAA	TAAGCATAGC	TGAGATTAAT	GATTCTCAGC	ATTAGTTATA	2820
158	AATAGAATAA	GCTGGAAAAA	CTTCACCTCC	CCTCCACCAC	CAGATCCCAA	AAACAAAATA	2880
160	CTGAAGATGC	TTATTTCAAT	ACTCAGGGAA	AATTTTCTTG	CCAAAAAGGC	AAGAATTGTA	2940
162	TAATTCATTC	ACTTATTTTA	TTTTTTTTTA	TTTTTAAGGT	CTAAGAGGAT	TTCAAAGTGA	3000

RAW SEQUENCE LISTING

DATE: 02/21/2002

PATENT APPLICATION: US/09/143,155

TIME: 09:52:14

Input Set : N:\Crf3\RULE60\09143155.txt

Output Set: N:\CRF3\02212002\I143155.raw

```

164 ATGCCCCCTC CTCAC TTTTG GTCAAATTGG AAATGGGGGT GAGATGAAGA GTTATAACAT 3060
166 ATAAC TAAAT GGACATTGTT CTCTATTCCA CAGAATTGAC TGCGACTGGA AATATGGCAA 3120
168 CTTTTCAATC CTTGCATCAT GCTACTAAGA TAATTTTAA ATGAGTATAC ATGGAACAAA 3180
170 AAATGAAACT TTATTCCTTT ATTTATTTTA TGCTTTTCA TCTTAATTTG AATTTGAGTC 3240
172 ATAAACCATA TACTTTCAAA ATGTTAATTC AACATTAGCA TAAAAGTTCA ATTTTAACTT 3300
174 GGAAATATCA TGAACATATC AAATTATGTA TAAAAATAAT TTCTGGAATT GTGATTATTA 3360
176 TTTCTTTAAG AATCTATTTT CTAACCAGTC ATTTCAATAA ATTAACCCTT AGGCATATTT 3420
178 AAGTTTTCTT GTCTTTATTA TATTTTAAA AATGAAATTG GTCTCTTTAT TGTAACTTA 3480
180 AATTTATCTT TGATGTTAAA AATAGCTGTG GAAAA 3515
182 (2) INFORMATION FOR SEQ ID NO: 2:
184 (i) SEQUENCE CHARACTERISTICS:
185 (A) LENGTH: 222 amino acids
186 (B) TYPE: amino acid
187 (D) TOPOLOGY: linear
189 (ii) MOLECULE TYPE: protein
191 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
193 Met Lys Val Leu Ile Leu Ala Cys Leu Val Ala Leu Ala Ile Ala Arg
194 1 5 10 15
196 Glu Gln Glu Glu Leu Asn Val Val Gly Glu Thr Val Glu Ser Leu Ser
197 20 25 30
199 Ser Ser Glu Glu Ser Ile Thr His Ile Asn Lys Lys Ile Glu Lys Phe
200 35 40 45
202 Gln Ser Glu Glu Gln Gln Gln Thr Glu Asp Glu Leu Gln Asp Lys Ile
203 50 55 60
205 His Pro Phe Ala Gln Ala Gln Ser Leu Val Tyr Pro Phe Thr Gly Pro
206 65 70 75 80
208 Ile Pro Asn Ser Leu Pro Gln Asn Ile Leu Pro Leu Thr Gln Thr Pro
209 85 90 95
211 Val Val Val Pro Pro Phe Leu Gln Pro Glu Ile Met Gly Val Pro Lys
212 100 105 110
214 Val Lys Glu Thr Met Val Pro Lys His Lys Glu Met Pro Phe Pro Lys
215 115 120 125
217 Tyr Pro Val Glu Pro Phe Thr Glu Ser Gln Ser Leu Thr Leu Thr Asp
218 130 135 140
220 Val Glu Lys Leu His Leu Pro Leu Pro Leu Val Gln Ser Trp Met His
221 145 150 155 160
223 Gln Pro Pro Gln Pro Leu Ser Pro Thr Val Met Phe Pro Pro Gln Ser
224 165 170 175
226 Val Leu Ser Leu Ser Gln Pro Lys Val Leu Pro Val Pro Gln Lys Ala
227 180 185 190
229 Val Pro Gln Arg Asp Met Pro Ile Gln Ala Phe Leu Leu Tyr Gln Glu
230 195 200 205
232 Pro Val Leu Gly Pro Val Arg Gly Pro Phe Pro Ile Leu Val
233 210 215 220

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/143,155

DATE: 02/21/2002

TIME: 09:52:15

Input Set : N:\Crf3\RULE60\09143155.txt

Output Set: N:\CRF3\02212002\I143155.raw

L:6 M:220 C: Keyword misspelled or invalid format, [(i) APPLICANT:]

L:27 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:28 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]